A phased-addition strategy enhances the yield of RNAs obtained by *in vitro* transcription with modified transcriptional initiators

Senior Honors Thesis

Presented in Partial Fulfillment for Graduating with a Honors Research Distinction in the College of Arts and Science of The Ohio State University

By
Katie Lynn Adib

The Ohio State University
2015

Examination Committee:

Venkat Gopalan, Advisor
Edward J. Behrman
Juan D. Alfonzo

Abstract

RNA is active in multiple biological processes, not only as an information carrier but in catalyzing reactions and regulating gene expression. To understand the basis for this versatility, it is critical to establish structure-function relationships in RNAs. Covalently-attached biochemical or biophysical probes (e.g., fluorophores) at site-specific positions are often used to uncover these correlates in target RNAs. Generation of such modified RNAs often requires the conjugation of molecular probes to suitable chemical functionalities that have been introduced at either terminal or internal positions in the target RNA. A simple approach to create a 5'-labeled RNA is through an *in vitro* transcription (IVT) using T7 RNA polymerase. Since IVTs of DNA templates containing a class III ϕ6.5 promoter by T7 RNA polymerase is preferentially initiated by GTP, guanosine analogs (nucleoside or nucleotide monophosphate) have been used as transcriptional initiators to introduce different reactive groups at the 5'-end of transcribed RNAs. Although adding guanosine analogs in a large molar excess over GTP at the start of transcription generates transcripts that are predominantly 5'-modified, the overall yield of RNA is seriously compromised due to the requirement for GTP during Based on the premise that the poor yield results from the rapid depletion of limited GTP, which was intentionally used at low concentrations to favor incorporation of the modified analog, we have investigated and validated a phased-addition strategy that results in high yields of full-length, 5'-modified RNAs from small-scale transcriptions. We examined this method using three guanosine analogs: 5'-deoxyguanosine-5'-monophosphorothioate

(GSMP), 5'-deoxy-5'-hydrazinylguanosine (NH_2NH-G), and 5'-azido-5'deoxyguanosine (N_3-G).

The phased-addition strategy is made up of two components that differ from a typical modified *in vitro* transcription. The first change is that a supplementation rather than a fixed-ratio approach is used. The former consists of supplementing GTP to the IVT during the first three or four hours of transcription as opposed to the fixed-ratio approach in which no additional GTP is added once transcription is initiated. The supplementation approach allows for a second change, which is to increase the initial guanosine analog:GTP ratio without concerns about the yield. Previously reported modified *in vitro* transcriptions using GSMP, NH_2NH-G, and N_3-G were typically preformed at the following ratios: 4:1 GSMP:GTP, ~31:1 NH_2NH-G:GTP, and 4:1 N_3-G:GTP. The supplementation of GTP permits the use of modified analog:GTP ratios as high as 50:1. These two adjustments to a typical *in vitro* transcription increase the total RNA yield to levels comparable to those from a typical *in vitro* transcription, while permitting ~90% incorporation of the modified guanosine analog at the 5'-end.

Dedication

This document is dedicated to my family, who made me the person I became. I will be eternally grateful for their love.

Acknowledgments

I thank my advisor Dr. Venkat Gopalan for his mentorship. He has been a great source of support and knowledge, and I appreciate his great patience along the way. I would also like to thank members of the Gopalan laboratory, in particular Dr. Lien B. Lai, Stella Lai, Tien-Hao Chen, and Ila Marathe for their help with experiments as well as their friendship.

I am grateful to Dr. Edward J. Behrman for his mentorship and for synthesizing the guanosine analogs. It has been wonderful learning the chemistry behind the compounds that I used, and I thank him for all his help in this endeavor. Additionally, I would like to express my gratitude to Dr. Jennifer J. Ottesen for use of the NanoDrop 2000 UV-Vis Spectrophotometer and to Dr. Dmitri Kudryashov for use of the TECAN microplate reader in their respective laboratories.

Vita

January 21, 1993…….. Born – Cincinnati, Ohio

2011 to present.……B.S., Biochemistry and Philosophy,
The Ohio State University

Fields of Study

Major Field: Biochemistry and Philosophy

Table of Contents

List of Figures

Abbreviations

ATP – 5'-adenosine-5'-triphosphate

ATPγS – 5'-*O*-(3-thio) adenosine triphosphate

Biotin-HPDP – N-[6-(Biotinamido)hexyl]-3'-(2'-pyridyldithio)propionamide

CTP – 5'-cytidine-5'-triphosphate

EDTA – ethylene diamine tetra acetate

GTP – guanosine-5'-triphosphate

GSMP – 5'-deoxyguanosine-5'-monophosphorothioate

IVT – *in vitro* transcription

N_3-G – 5'-azido-5'deoxyguanosine

NH_2NH-G – 5'-deoxy-5'-hydrazinylguanosine

NMR – nuclear magnetic resonance

Nt – nucleotides

rNTP – ribose nucleoside 5'-triphosphate

PCR – polymerase chain reaction

pre-tRNA – precursor-transfer RNA

T4 PNK – T4 polynucleotide kinase

UTP – uridine-5'-triphosphate

Chapter 1
Introduction

1.1 Biological versatility of RNA

RNA is a macromolecule found in all known forms of life. While RNA initially was believed to be a simple information carrier, during the last three decades we have begun to appreciate that RNA is active in multiple biological processes, including catalyzing reactions, fighting infectious agents, and regulating gene expression (Cech, 2012; Guerrier-Takada et al., 1983; Kruger et al., 1982). The biological versatility of RNA has elicited interest in making it a target for new drugs and in exploring it as a tool for novel therapies (Blakeley et al., 2012). RNAs are linked to various diseases, making it a target for small molecules that can modulate function. The different types of biological functions of RNA correlate with its structural plasticity. Covalently-attached biochemical or biophysical probes (e.g., fluorophores, paramagnetic probes) at specific positions can be used to uncover structure-function relationships in target RNAs. Generation of such modified RNAs, however, often requires the conjugation of molecular probes to suitable chemical functionalities that have been introduced at either terminal or internal positions in the target RNA (Pagano et. al., 2011; Paredas et al., 2011; Qin and Pyle, 1999; Zhang et al., 2001).

1.2 Methods to label RNA

Whether it is through enzymatic methods or solid-phase synthesis, there are multiple approaches to label an RNA. Each method has its benefits, as well as its own

shortcomings. Solid-phase synthesis of RNA allows for direct modification as well as a variety of post-synthetic modifications compared to enzymatic methods (Paredas et al., 2011). Solid-phase synthesis allows for terminal and internal labeling. However, there are some limitations when using synthetic RNA to introduce a modification. Foremost, modifications need to be incorporated as phosphoramidites at the specified location (Paredas et al., 2011; Qin and Pyle, 1999). Another limitation pertains to the length of RNA synthesized, typically 40 nucleotides or less (Qin and Pyle, 1999).

For longer RNAs, post-transcriptional methods or enzymatic routes can be utilized. For example, following *in vitro* transcription, a 3'-terminal modification can be effected though periodate oxidation (Qin and Pyle, 1999). Periodate oxidizes the 3'-terminal diol to generate a dialdehyde, introducing a chemical functionality that renders it reactive towards a variety of compounds including hydrazine derivatives. Enzymatic RNA synthesis, similar to solid-phase synthesis, can be used to introduce modifications at either terminus or at an internal location, although the latter entails an additional ligation step. For example, T4 polynucleotide kinase (T4 PNK), can introduce a 5'-thiol following RNA synthesis when it transfers the γ-phosphorothioate of ATPγS to the 5'-hydroxyl of the target RNA (Pagano et al., 2011). Introducing this unnatural chemical functionality enables reaction with reagents such as 5-iodoacetamidofluorescein (5-IAF; Pagano et al., 2011).

It is important to appreciate that there are methods to attach a label on the RNA without the need for a modification post-synthesis. Through the use of a DNA template with an unnatural base (introduced through solid-phase DNA synthesis), an unnatural

base can be introduced during *in vitro* transcription of an RNA (Domnick et al., 2015; Tor and Dervan, 1993). In the recent example cited, the yield of full-length transcripts was reported to be low; 36% of the transcripts were full length when using DNA with one unnatural base, and 16% when using DNA with two unnatural bases (Domnick et al., 2015). Another approach to introduce a label to the RNA is through the hybridization of a complementary fluorescent-labeled DNA or PNA oligonucleotide (Schmitz et al., 2014). A caveat to this method is that the hybridization might alter the structure and therefore the function of the target RNA.

1.3 Statement of the problem

A simple approach to introduce a new chemical functionality in an RNA is accomplished by introducing a 5'-modification through use of a T7 RNA polymerase-based *in vitro* transcriptions (IVTs) of DNA templates. Depending on the promoter used, guanosine or adenosine analogs can be used to introduce a 5'-modification (Paredas et al., 2011). The polymerase will incorporate a guanosine or adenosine analog at the transcriptional initiation site instead of GTP or ATP, illustrating the enzyme's tolerance for incorporating modified nucleotide analogs.

Although there are multiple approaches to introduce new chemical functionalities into an RNA, my study focuses on a 5'-modification introduced during IVTs of DNA templates containing a class III ϕ6.5 promoter by T7 RNA polymerase. With this promoter, GTP is used preferentially to initiate transcription; therefore, guanosine analogs

(nucleoside or nucleotide monophosphate) have been used as transcriptional initiators to introduce different reactive groups at the 5'-end of transcribed RNAs.

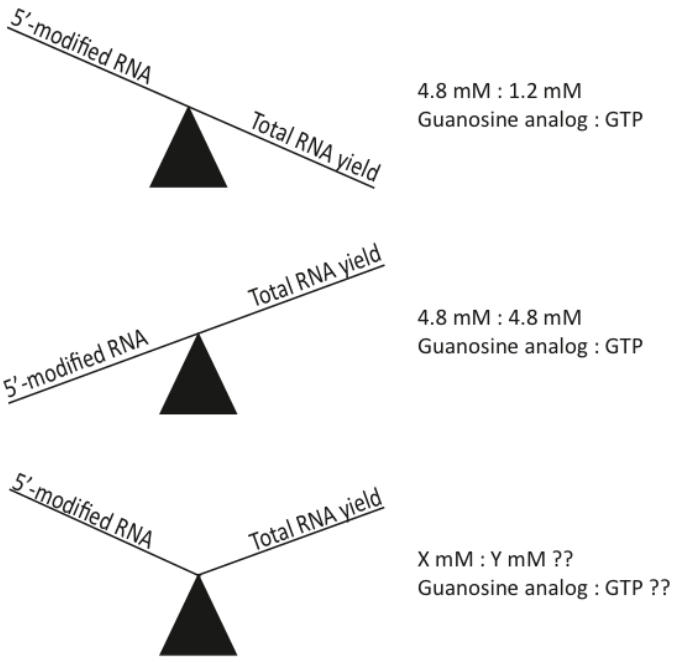
The three guanosine analogs used in this study were synthesized by Dr. Edward Behrman (OSU): 5'-deoxyguanosine-5'-monophosphorothioate (GSMP), 5'-deoxy-5'-hydrazinylguanosine (NH_2NH-G), and 5'-azido-5'-deoxyguanosine (N_3-G) (Figure 1). The following reasoning inspired the choice of these three guanosine analogs. GSMP was used because a thiol would be generated with an alkaline phosphatase treatment post-transcription. Thiols react with many compounds, including iodoacetamides, maleimides, and bromomethyl ketones, allowing for a wide array of different biochemical or biophysical probes to be introduced into the thiol-containing target RNA (Zhang et al., 2001). NH_2NH-G was studied, because hydrazines are more reactive than amines and will react with active esters (e.g., commonly available NHS-esters) and aldehydes (Raddatz et al., 2002; Skipsey et al., 2013). N_3-G was studied because of the recent wide use of click chemistry, one example being the copper(I)-promoted azide-alkyne cycloaddition reaction (Paredas and Das, 2011).

These guanosine analogs can only be used as the initiator nucleotide because they cannot promote formation of a phosphodiester bond, a key attribute of elongation (Figure 2). During the elongation steps in transcription, the release of pyrophosphate from the incoming rNTP provides the driving force for bond formation. Therefore, these guanosine analogs compete with GTP only during initiation. Due to this reason, GTP is preferentially depleted compared to the guanosine analog.

Adding guanosine analogs in a large molar excess over GTP at the start of transcription generates transcripts that are predominantly 5'-modified, but the overall yield of RNA is greatly compromised (Scheme 1). However, if the guanosine analog and GTP were added in equal molar amounts, the overall yield of RNA will be high, but the amount of labeled RNA will decrease (Scheme 1). Andrew Wallace, a former undergraduate and subsequently a graduate student in the Gopalan laboratory, reported that using equimolar amounts of the guanosine analog and GTP in an IVT resulted in only 37% of the RNA transcripts being 5'-modified (Wallace, 2013). While a higher amount of guanosine analog to GTP might render it more favorable for initiation by T7 RNA polymerase, an excess of guanosine analog over GTP would be expected to cause premature termination of transcription due to the rapid depletion of GTP. I set out to evaluate a method in which neither the overall yield of RNA or the 5'-modified RNA are compromised (Scheme 1).

1.4: Solution to the problem: a phased-addition strategy

An assay to determine the yield of modified RNA from an *in vitro* transcription using a 4:1 ratio of N_3-G:GTP showed that 87% of the transcripts were initiated with N_3-G; these studies were performed with a tetramer model RNA (Wallace, 2013). While the initial 4:1 ratio of N_3-G:GTP might lead to the expectation that only 80% of the transcripts should be modified, it is in fact not the case due to the rapid depletion of GTP. As GTP is consumed during an IVT, the guanosine analog:GTP ratio is probably



Scheme 1: Two different scenarios are depicted to illustrate the problem of concomitantly achieving high 5'-modification and high RNA yield during *in vitro* transcription. (Top) A high concentration of the guanosine analog leads to a high yield of modified RNA but low total RNA yield. (Middle) However, equal concentration of the guanosine analog and GTP leads to low 5'-modification but high total RNA yield. (Bottom) The overall goal of this research is to design a method that will lead to mostly modified RNA at high RNA yield.

greater than 4:1. The idea that the ratio is not constant between the guanosine analog and GTP during *in vitro* transcription is the basis for the phased-addition strategy, which was developed to ensure that neither the total yield nor the yield of modified RNA is compromised during transcription.

The phased-addition strategy differs from a normal transcription in two ways. The first is through the introduction of a supplementation approach (Figure 3). During the transcription, the reaction is supplemented with defined amounts of GTP. Supplementation with GTP allows full-length transcripts to still be made during multiple sequential rounds of IVT, even while ensuring that GTP will not out-compete the

guanosine analog at initiation. Due to the supplementation approach, I was able to implement the second component of the phased-addition strategy. Since rapid depletion of GTP is no longer a concern, the initial guanosine analog:GTP ratio could be increased to levels never tested before.

I investigated and validated a phased-addition strategy that results in high yields of full-length, 5'-modified RNAs from small-scale transcriptions and used evidence from spectrophotometric assays to accurately calculate of the incorporation efficiency of each guanosine analog. Results from experiments with three guanosine analogs (GSMP, NH_2NH-G, and N_3-G) demonstrate that the phased-addition strategy enhances the yield of RNA obtained by *in vitro* transcription while improving the extent of modification.

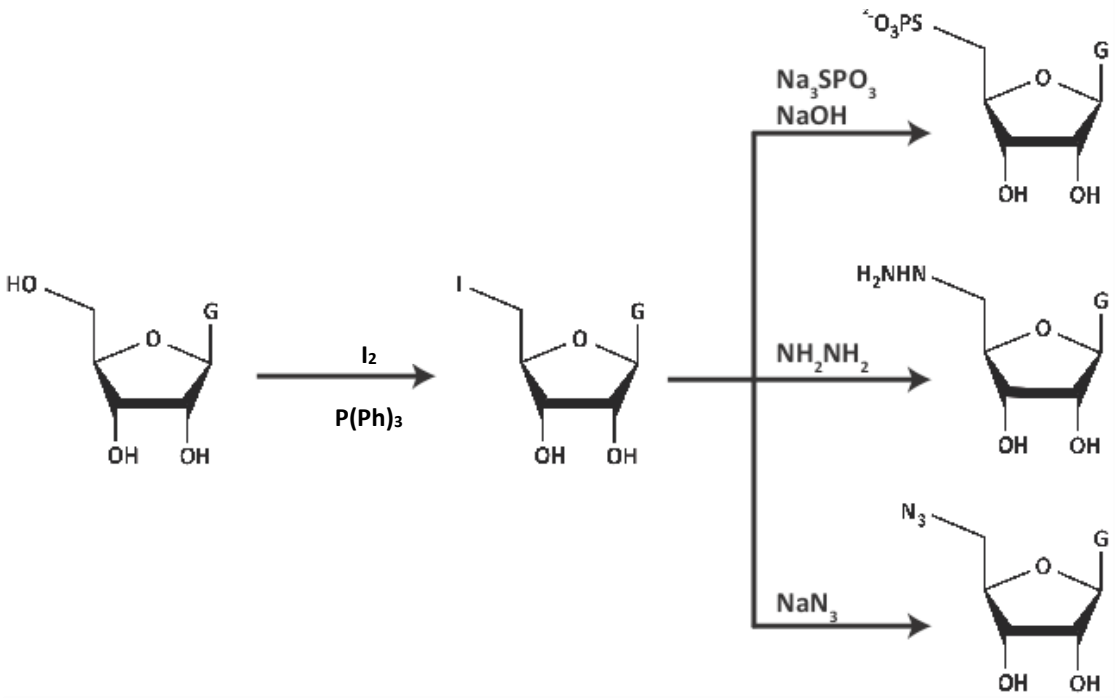


Figure 1. Synthesis of three guanosine analogs from 5'-iodo-5'-deoxyguanosine. Synthesis of 5'-deoxyguanosine-5'-monophosphorothioate, 5'-deoxy-5'-hydrazinyl guanosine, and 5'-azido-5'deoxyguanosine is depicted.

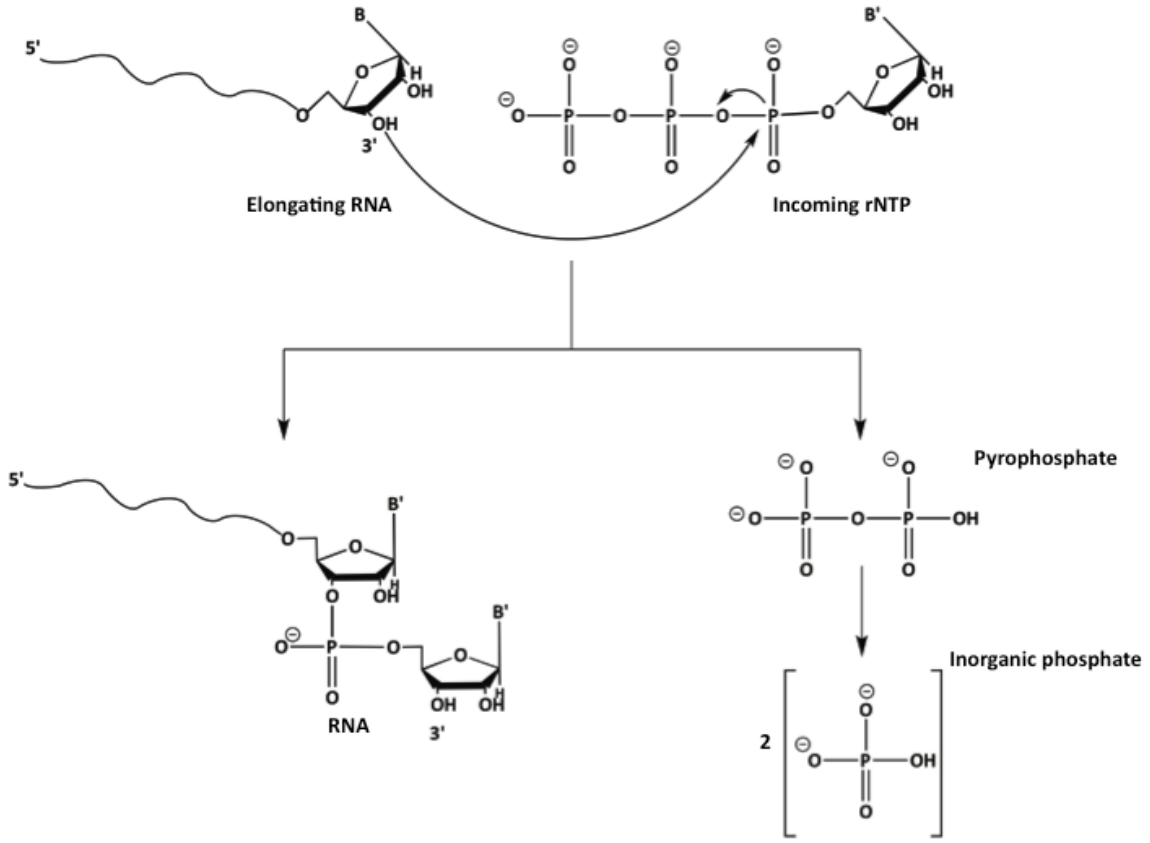


Figure 2. Phosphodiester bond formation during RNA elongation by an RNA polymerase. Release of pyrophosphate provides the energy to allow incorporation of the incoming nucleotide into the growing RNA chain.

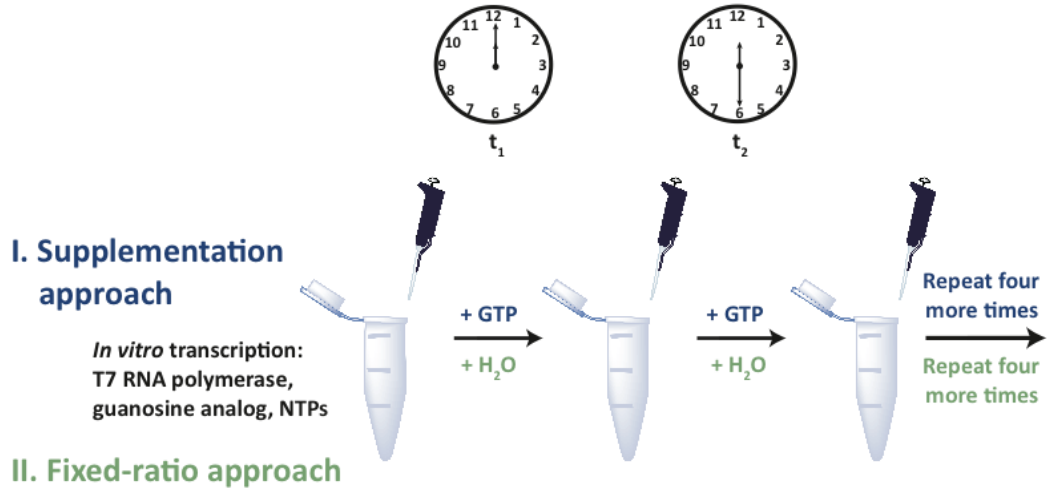


Figure 3. Phased-addition strategy. This strategy entails the addition of GTP every 30 min for the first three or four h of a modified *in vitro* transcription. The fixed-ratio approach where the initial guanosine analog:GTP ratio is held constant has been typically used for modified *in vitro* transcription. The supplementation approach allows for an increase in the initial guanosine analog:GTP ratio without concerns about the rapid depletion of GTP during elongation.

Chapter 2
Materials and Methods

2.1 Preparation of DNA template

For all the studies reported here, an *Arabidopsis thaliana* pre-tRNA^{Cys} (149 nt) was used as the model. This pre-tRNA was transcribed using a template generated by PCR. The PCR amplification employed Q5® polymerase (New England BioLabs® Inc.), the universal forward primer (5'-CGACGTTGTAAAACGACGGCCAG-3'), the universal reverse primer (5'-GGAAACAGCTATGACCATGAT-3'), and pBT7-pre-tRNA^{Cys} as the template.

2.2: Standard *in vitro* transcription

pre-tRNA^{Cys} was generated through a run-off *in vitro* transcription of PCR-generated linear templates (as described above). The reactions were preformed for 16 h at 37°C in 80 mM HEPES (pH 7.5), 33 mM $MgCl_2$, 1 mM spermidine, 10 mM dithiothreitol, 6 mM rNTPs (ATP, CTP, GTP, UTP), and in-house purified T7 RNA polymerase. The RNA was purified by sequential DNase I treatment at 37°C for 1 h, phenol-chloroform extraction, dialysis to remove unincorporated rNTPs, and ethanol precipitation. This purification method was employed for all RNAs used in this study. The yield of RNA was determined by measuring the absorbance at 260 nm.

2.3 Incorporation of 5'-deoxyguanosine-5'-monophosphorothioate, 5'-deoxy-5'-hydrazinylguanosine, 5'-azido-5'-deoxyguanosine into pre-tRNA^{Cys}

pre-tRNA^{Cys} was generated through a run-off *in vitro* transcription as described above with one key modification. While ATP, CTP, and UTP were 6 mM, the modified

guanosine analog and GTP were used at ratios specified in the text. For the phased-addition strategy, either 0.5 mM or 1 mM GTP (as specified) was added every 30 min for three or four h (Figure 4). The RNA was purified and the yield quantitated in an identical fashion to the standard transcriptions.

Following transcription with GSMP, the RNA transcripts were subjected to treatment with calf intestinal phosphatase (0.5 units/µg RNA) at 37°C for one h. The RNA was purified by phenol-chloroform extraction and ethanol precipitation. The resulting RNA should have a 5'-thiol if the transcript was indeed initiated with GSMP.

2.4 Quantitation of 5'-sulfhydryl incorporation in an RNA

Ellman's test was used to determine the amount of 5'-thiol in the RNA (Ellman, 1959). Ellman's reagent [5,5'-dithio-bis-(2-nitrobenzoic acid)] was mixed with known concentrations of L-cysteine (0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55 or 0.60 mM) in 0.1 M sodium phosphate (pH 8), 1 mM EDTA and incubated for 15 min at ~24°C. A NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific) was then used to measure the absorbance at 412 nm, and generate a standard curve. By using this reference plot, the amount of 5'-thiol incorporation in any RNA could be determined.

2.5 Quantitation of 5'-hydrazinyl incorporation in an RNA

To quantitate the amount of transcripts initiated with NH_2NH-G, I adopted a spectrophotometric test typically used to measure the amount of hydrazine in the air (Kaveeshwar and Gupta, 1992). NH_2NH-G was standardized through a potassium

permanganate titration, and then known concentrations of NH_2NH-G (0.006, 0.012, 0.019, 0.024, 0.028, 0.033, 0.038, 0.043, 0.047 or 0.052 mM) were reacted with veratraldehyde [2% (v/v) in 25% (v/v) ethanol] at 55°C for one h. The resulting mixture was then treated with 2.5 M sulfuric acid and the absorbance of the hydrazone product was measured at 410 nm using a NanoDrop 2000 UV-Vis Spectrophotometer. A linear colorimetric plot was established and used to quantitate the incorporation of NH_2NH-G in an RNA.

2.6 Quantitation of 5'-azido incorporation in an RNA

To quantitate the amount of transcripts initiated with N_3-G, I adopted a spectrophotometric test typically used to measure the amount of azide in the air or soda (Christova-Bagdassarian and Atanassova, 2007; Tsuge et al., 2001). Known concentrations of N_3-G (0.03, 0.06, 0.15, 0.30, 0.45, 0.60, 0.75, 0.91, 1.06, 1.21, 1.36, 1.51 or 1.66 mM) were reacted with 4 M hydrochloric acid and 0.6 M ferric chloride. The reactions were mixed at ~24°C and the absorbance of the ferric azide complex could be immediately measured at 460 nm using a NanoDrop 2000 UV-Vis Spectrophotometer. Using this colorimetric plot, the incorporation of N_3-G in any RNA could be quantitated.

2.7 Determination of the extinction coefficient for 5'-deoxyguanosine-5'-monophosphorothioate, 5'-deoxy-5'-hydrazinylguanosine, 5'-azido-5'-deoxyguanosine.

To determine the extinction coefficient of each guanosine analog, known concentrations of each guanosine analog were made (25, 50, 75, and 100 μM) in phosphate buffer (pH 7). The absorbance was then measured at 251 nm, which is the maximum wavelength for absorbance for guanosine. The absorbance values were plotted against the concentration, and each extinction coefficient was determined from the slope of the linear plot generated.

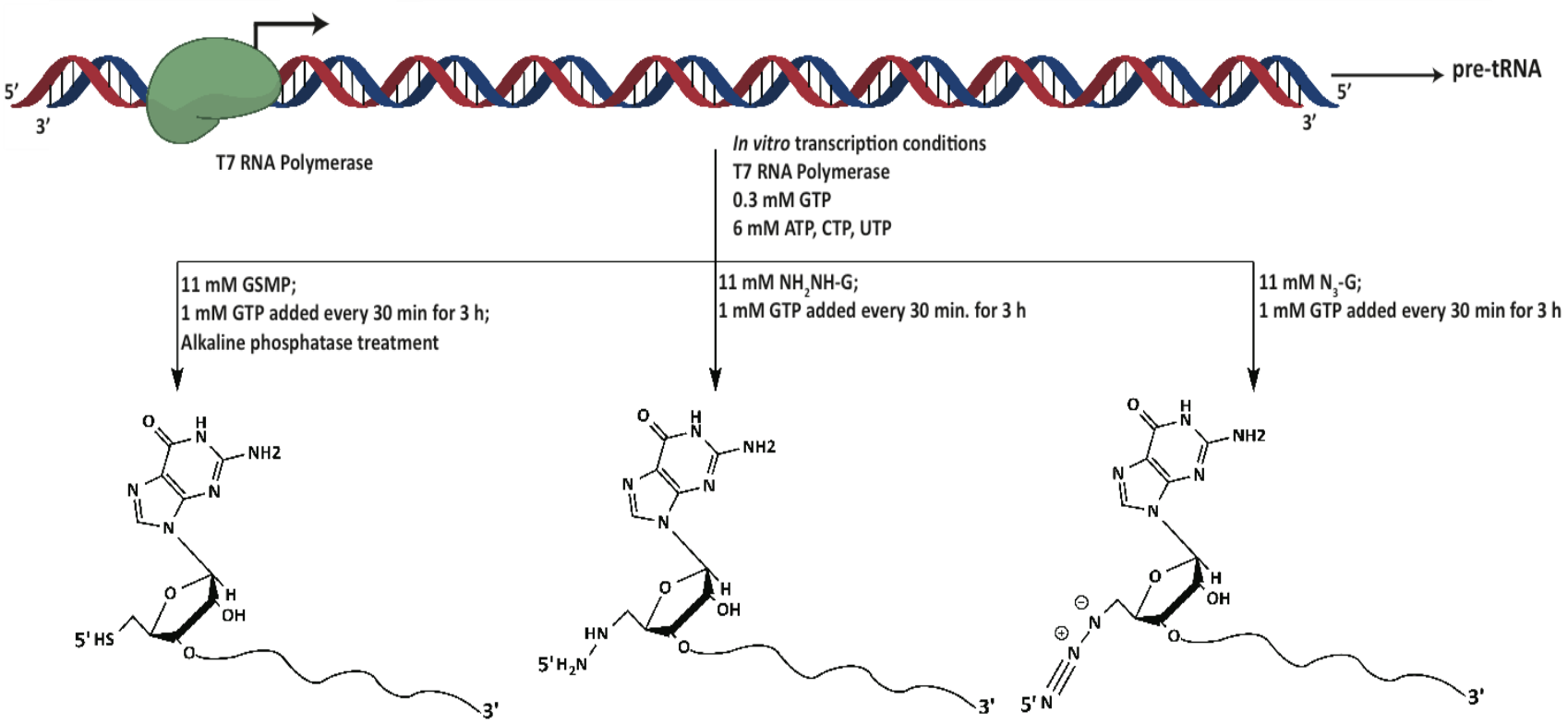


Figure 4. Phased-addition strategy to enhance initiation with GSMP, NH_2NH-G, and N_3-G during an IVT.

Chapter 3
Results

3.1 Extinction coefficient of 5'-deoxyguanosine-5'-monophosphorothioate, 5'-deoxy-5'-hydrazinylguanosine, 5'-azido-5'-deoxyguanosine

The extinction coefficient of each guanosine analog was determined. For GSMP, NH_2NH-G, and N_3-G, the extinction coefficients were determined to be 13,693, 13,208, and 13,145 $M^{-1}\,cm^{-1}$, respectively (Figure 5). These values are comparable to that reported for guanosine 5'-monophosphate (13,650 $M^{-1}\,cm^{-1}$; Behrman, 2000).

3.2 Incorporation of 5'-deoxyguanosine-5'-monophosphorothioate

5'-Deoxyguanosine-5'-monophosphorothioate was synthesized by reacting 5'-iodo-5'-deoxyguanosine with thiophosphate in sodium hydroxide. Selective precipitation using methanol yielded GSMP of ~65% purity (Brear et al., 2007).

There are reports of successful incorporation of GSMP in IVTs of short RNAs using T7 RNA polymerase and a 4:1 GSMP:GTP ratio. For example, Zhang et al. (2001) reported 60% 5'-incorporation, which was determined by reacting the thiol with biotin-HPDP. The transcripts with a 5'-thiol were biotinylated with the biotin-HPDP and subsequently could be detected as band-shifts through the use of a streptavidin-based gel-shift assay. By using a phosphorimager**,** the fraction of the RNA with a 5'-modification relative to the total RNA was quantitated by analyzing the intensity of the individual bands in the gel-shift assay (Zhang et al., 2000). In my studies, using a 4:1 GSMP:GTP ratio, the RNA yield was 67 ± 14 µg from a 100 µL IVT and 82 ± 1% of the transcripts had been initiated with GSMP (Figure 6). The incorporation of GSMP was determined through the use of Ellman's

test, which is typically used for quantitation of protein thiols (Ellman, 1959) (Figure 7). In a typical *in vitro* transcription, where there is no modified guanosine analog, the yield was 137 ± 8 µg from a 100 µL IVT, whereas the yield from a transcription using 4:1 GSMP:GTP ratio was only 67 µg from a 100 µL IVT (Figure 6). Each set of results was obtained from three independent trials.

To evaluate the phased-addition strategy, my initial tests were performed using a 11:1 GSMP:GTP ratio with subsequent additions of 0.5 mM GTP every 30 min for the first three h of the IVT. I determined from three independent trials that the yield was 104 ± 2 µg from a 100 µL IVT, representing a 1.5-fold increase from the 67 µg obtained using a 4:1 GSMP:GTP ratio IVT (Figure 6). Additionally, the incorporation of 5'-thiol was found to be 91 ± 1% (Figure 6). However, to conclusively prove that the phased-addition approach led to these remarkable gains, I performed three independent *in vitro* transcriptions using a 11:1 GSMP:GTP ratio, but without any supplementation of GTP during the IVT. Using this approach, the total yield of RNA was 23 ± 3 µg from a 100 µL IVT, with 92 ± 0.5% of the transcripts being initiated with GSMP (Figure 6).

To ensure that the supplementation of GTP for only the first three hours was sufficient, I extended the phased-addition to four hours. A single trial revealed that the yield was 107 µg from a 100 µL IVT, confirming that RNA yield (107 versus 104 µg) does not change significantly by extending the supplementation from three to four hours.

While these results were an improvement over those reported (Zhang et al., 2000; Zhang et al., 2001), I sought to determine if further changes to the initial GSMP:GTP ratio might afford additional gains. I increased the GSMP:GTP ratio to 37:1, and increased the

GTP supplementation from 0.5 mM to 1 mM every 30 min for three h. The RNA yield increased to 122 ± 2 μg from a 100 μL IVT and 5'-thiol incorporation was 90 ± 1% (Figure 6). With these new conditions, the yield increased 2-fold compared to the fixed-ratio approach using a 4:1 GSMP:GTP ratio.

3.3 Incorporation of 5'-deoxy-5'-hydrazinylguanosine

5'-Deoxy-5'-hydrazinylguanosine was synthesized by reacting 5'-iodo-5'-deoxyguanosine with hydrazine. The resulting compound was purified by precipitation with methanol and recrystallization from methanol (Brear et al., 2007). ^{1}H NMR spectrum of NH_2NH-G indicates that the purity is ~80% (Figure 8).

Successful incorporation of NH_2NH-G in an RNA generated by T7 RNA polymerase has been reported (Skipsey et al., 2013). In this previous study, a 31.3:1 NH_2NH-G:GTP ratio was employed. Approximately 30-40% incorporation of NH_2NH-G was estimated although with a low yield (numbers not reported, however). The incorporation of NH_2NH-G was estimated by measuring the amount of fluorescence after reacting the NH_2NH-G-containing RNA with fluorescien isothiocyanate.

For a direct measure of the incorporation of NH_2NH-G during an *in vitro* transcription, I sought a spectrophotometric assay that would measure the amount of hydrazine for NH_2NH-G either free or incorporated in an RNA. Although I did not find reports of such an assay, a spectrophotometric test for determining hydrazine in the air based on a reaction with veratraldehyde has been documented (Kaveeshwar and Gupta, 1992). I adopted this assay for my objective. I used NH_2NH-G and first established a linear

colorimetric response (Figure 9). With this standard reference curve in hand, I could determine the amount of NH_2NH-G incorporated in an RNA.

Consistent with the findings of Skipsey et al. (2013), a 31:1 NH_2NH-G:GTP ratio resulted in a yield of 10 µg of pre-$tRNA^{Cys}$ from a 100 µL IVT and 37% of the transcripts were initiated with NH_2NH-G (Figure 10). However, when I used an initial 11:1 NH_2NH-G:GTP ratio coupled with additions of 0.5 mM GTP every 30 min for three h, the RNA yield increased to 78 ± 3 µg from a 100 µL IVT. The NH_2NH-G incorporation was determined to be 26% (Figure 10).

To increase the extent of incorporation of NH_2NH-G, the NH_2NH-G:GTP ratio was increased to 22:1 with additions of 0.5 mM GTP every 30 min for three h. Although the yield of RNA decreased from 78 µg to 44 µg, the NH_2NH-G incorporation increased from 26% to 42%. In fact, with a 37:1 NH_2NH-G:GTP ratio, coupled with additions of 1 mM GTP every 30 min for the first three hours of transcription, the RNA yield improved further to 104 µg from a 100 µL IVT with 55% of the transcripts being initiated with NH_2NH-G (Figure 10).

Further investigation with a 40:1 NH_2NH-G:GTP or a 50:1 NH_2NH-G:GTP ratio coupled with 1 mM GTP addition every 30 min for four h showed further improvements with respect to both RNA yield and incorporation of NH_2NH-G. The 40:1 NH_2NH-G:GTP ratio resulted in 116 µg of pre-$tRNA^{Cys}$ from a 100 µL IVT with 58% NH_2NH-G incorporation, and the 50:1 NH_2NH-G:GTP ratio generated 105 µg with 62% incorporation (Figure 10). These results are from a single trial, and need to be repeated.

3.4 Incorporation of 5'-azido-5'-deoxyguanosine

5'-Azido-5'-deoxyguanosine was synthesized by reacting 5'-iodo-5'-deoxyguanosine with sodium azide. The resulting compound was purified through crystallization by cooling the reaction mixture (Brear et al., 2007).

There are reports of successful incorporation of N_3-G using *in vitro* transcriptions with T7 RNA polymerase and a 4:1 ratio of N_3-G:GTP. For example, Paredas et al. (2011) determined that a 4:1 N_3-G:GTP ratio results in almost complete modification at the 5'-end. The incorporation extent was determined by assessing a copper (I) mediated click ligation between an RNA with a 3'-alkyne functional group and an RNA with a 5'-azido group. Following ligation, ethidium bromide staining of a polyacrylamide gel was used to determine the extent of unreacted transcripts. Paredas et al. (2011) concluded that N_3-G almost completely replaces GTP at the initiation site due to the low number of unreacted transcripts, although quantitation was not provided (Paredas et al., 2011). My studies showed that a 4:1 N_3-G:GTP ratio resulted in a yield of 69 ± 8 µg of pre-tRNA^{Cys} from a 100 µL IVT with 84 ± 0.4% of the transcripts being initiated with N_3-G (Figure 11).

The incorporation of N_3-G was determined through a spectrophotometric test, adopted from an assay used to measure free azide in the air and in soda (Christova-Bagdassarian and Atanassova, 2007; Tsuge et al., 2001). The assay is based on the production of a ferric-azide chromophore resulting from the reaction of free azide with ferric chloride (Figure 12a). Although this was not an assay developed for determination of N_3-G, either free or incorporated into an RNA, the use of N_3-G instead of free azide

revealed a linear colorimetric response (Figure 12b). With this standard reference curve in hand, I could determine the amount of N_3-G incorporated in an RNA.

To improve the extent of incorporation of N_3-G, I preformed an *in vitro* transcription at a 37:1 N_3-G:GTP and added 1 mM GTP every 30 min for the first three h of transcription. The pre-$tRNA^{Cys}$ yield was 116 ± 2 μg from a 100 μL IVT with 92 ± 2% of the transcripts initiated by N_3-G (Figure 11). These results represent a 1.7-fold increase in the RNA yield compared to a typical 4:1 N_3-G:GTP ratio.

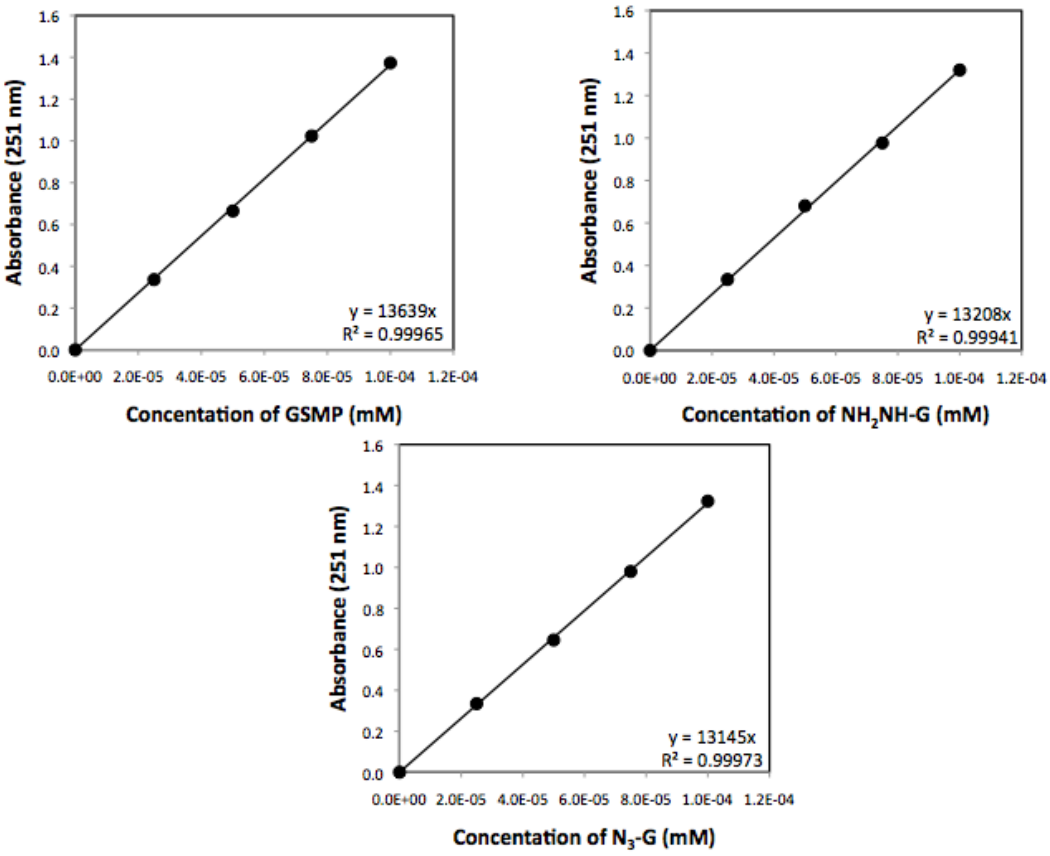


Figure 5: Extinction coefficients of GSMP, NH_2NH-G, and N_3-G.

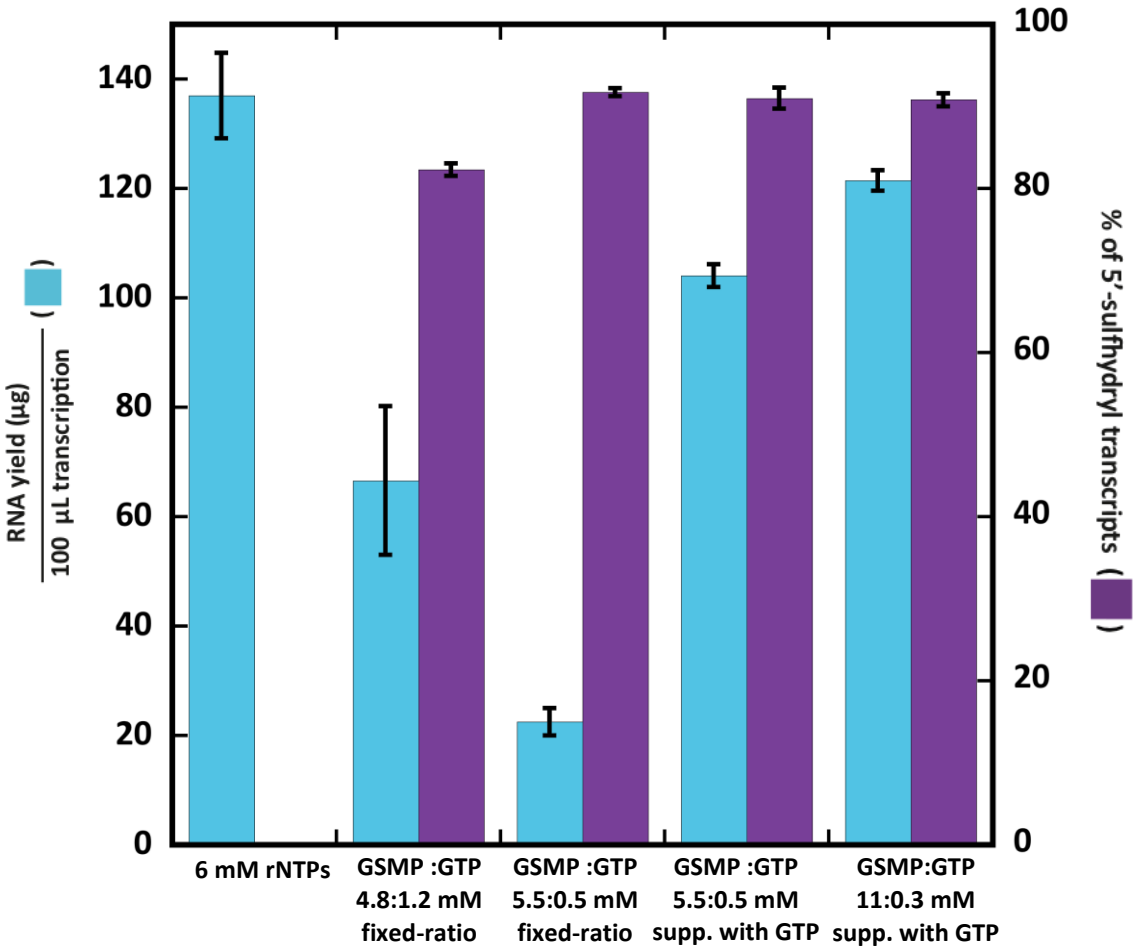


Figure 6. IVT using GSMP as the transcriptional initiator: results from the fixed-ratio and phased-addition approaches. The blue bars represent the yield of RNA in µg from a 100 µL IVT and the purple bars represent the percent of transcripts with a 5'-sulfhydryl substitution.

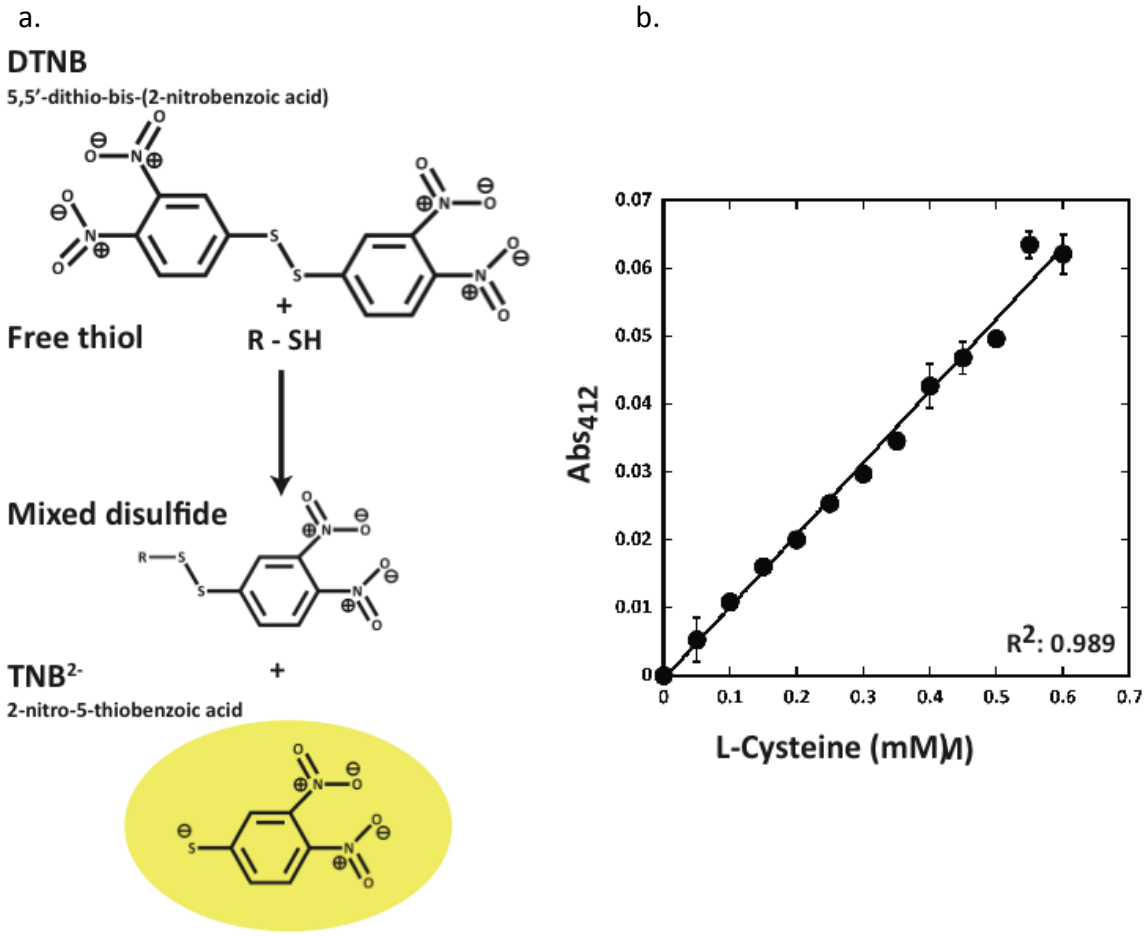


Figure 7. Quantitation of 5'-sulfhydryl incorporation in an RNA. (a) Ellman's test: reaction of DTNB [5,5'-dithio-bis-(2-nitrobenzoic acid)] with a free thiol produces TNB^{2-} (2-nitro-5-thiobenzoic acid), which absorbs light at 412 nm. (b) Linear colorimetric standard curve with L-cysteine.

a.

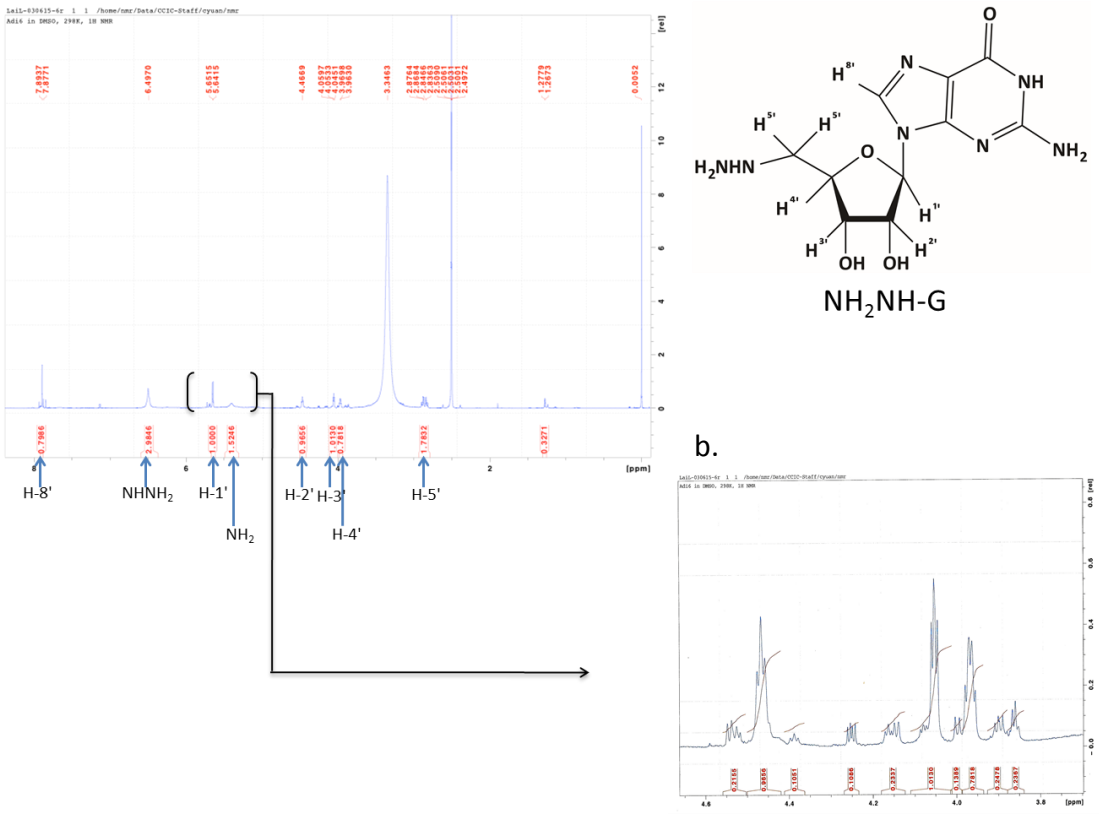


Figure 8. (a) ^{1}H NMR of NH_2NH-G. (b) Magnification of the spectrum from 3.8 to 4.6 ppm to highlight peaks corresponding to contaminants in the NH_2NH-G preparation.

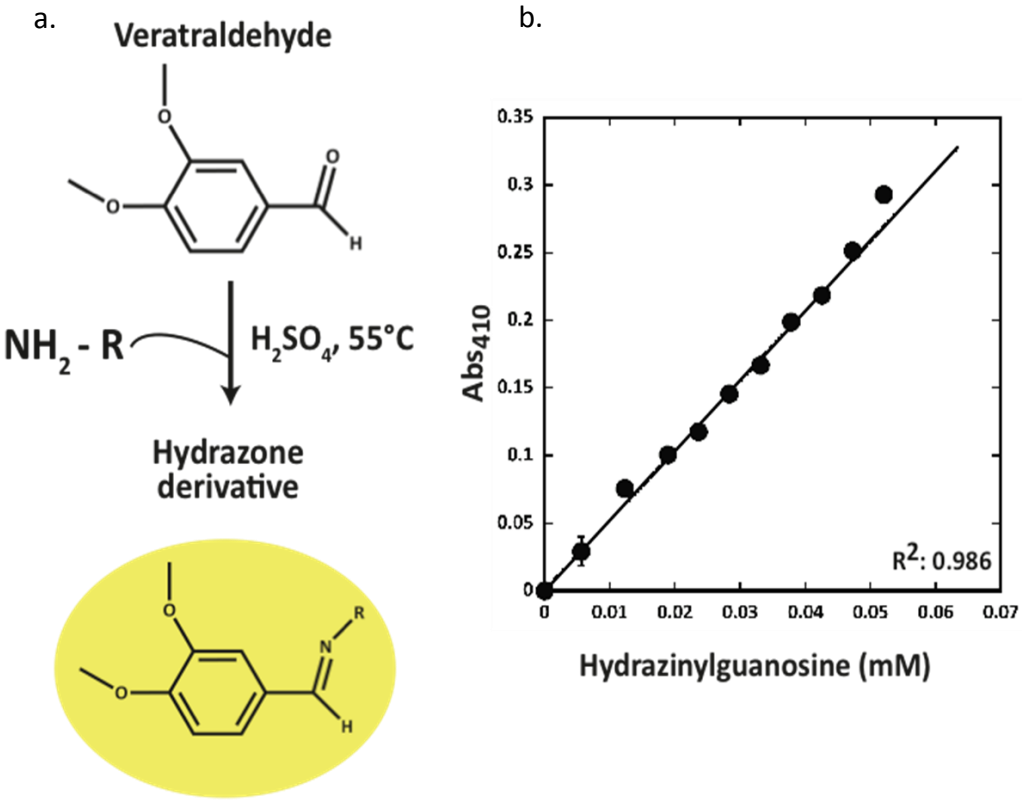


Figure 9. Quantitation of 5'-hydrazinyl incorporation in an RNA. (a) Reaction of veratraldehyde with a hydrazine to produce a hydrazone derivative that absorbs light at 410 nm. (b) Linear colorimetric standard curve with NH_2NH-G.

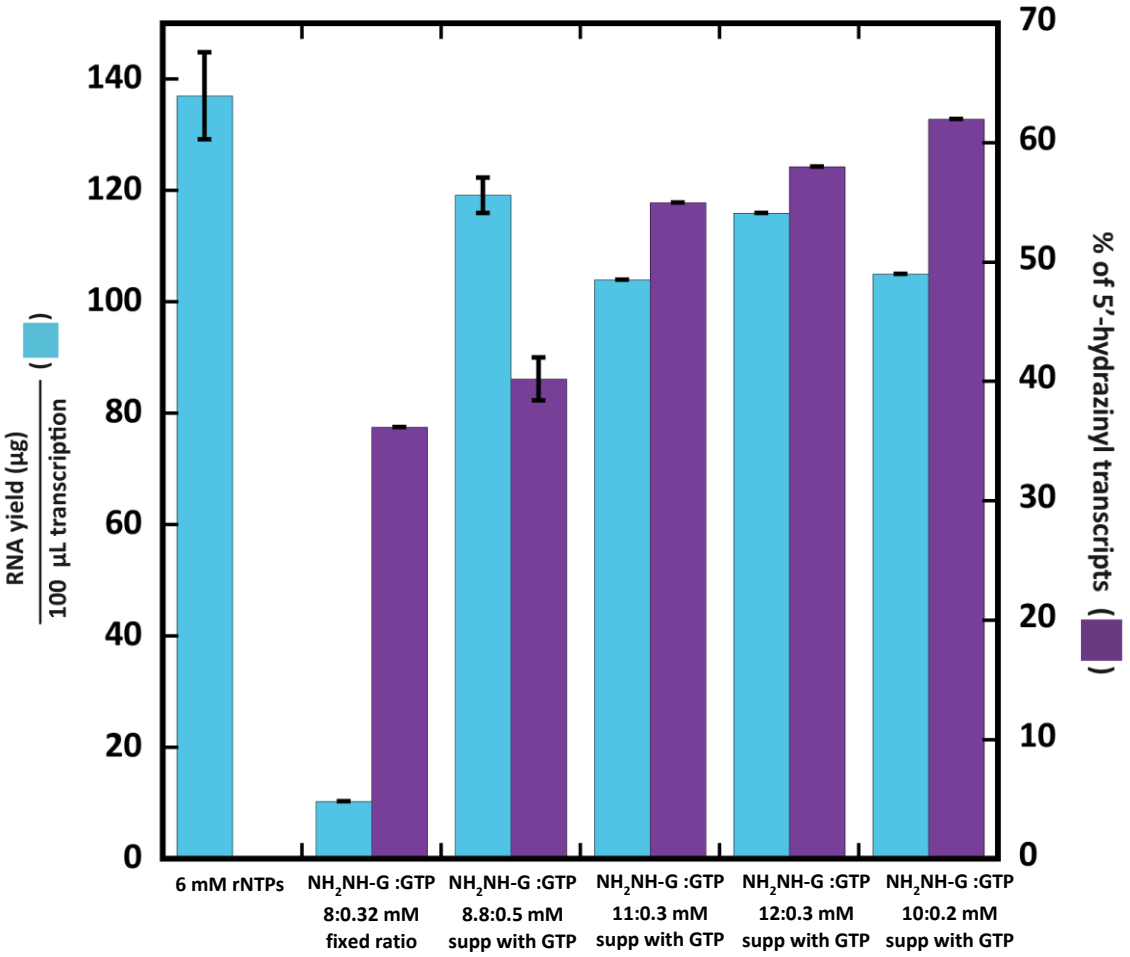


Figure 10. IVT using NH_2NH-G as the transcriptional initiator: results from fixed-ratio and phased-addition approaches. The blue bars represent the yield of RNA in µg from a 100 µL IVT and the purple bars represent the percent of transcripts with a 5'-hydrazinyl substitution.

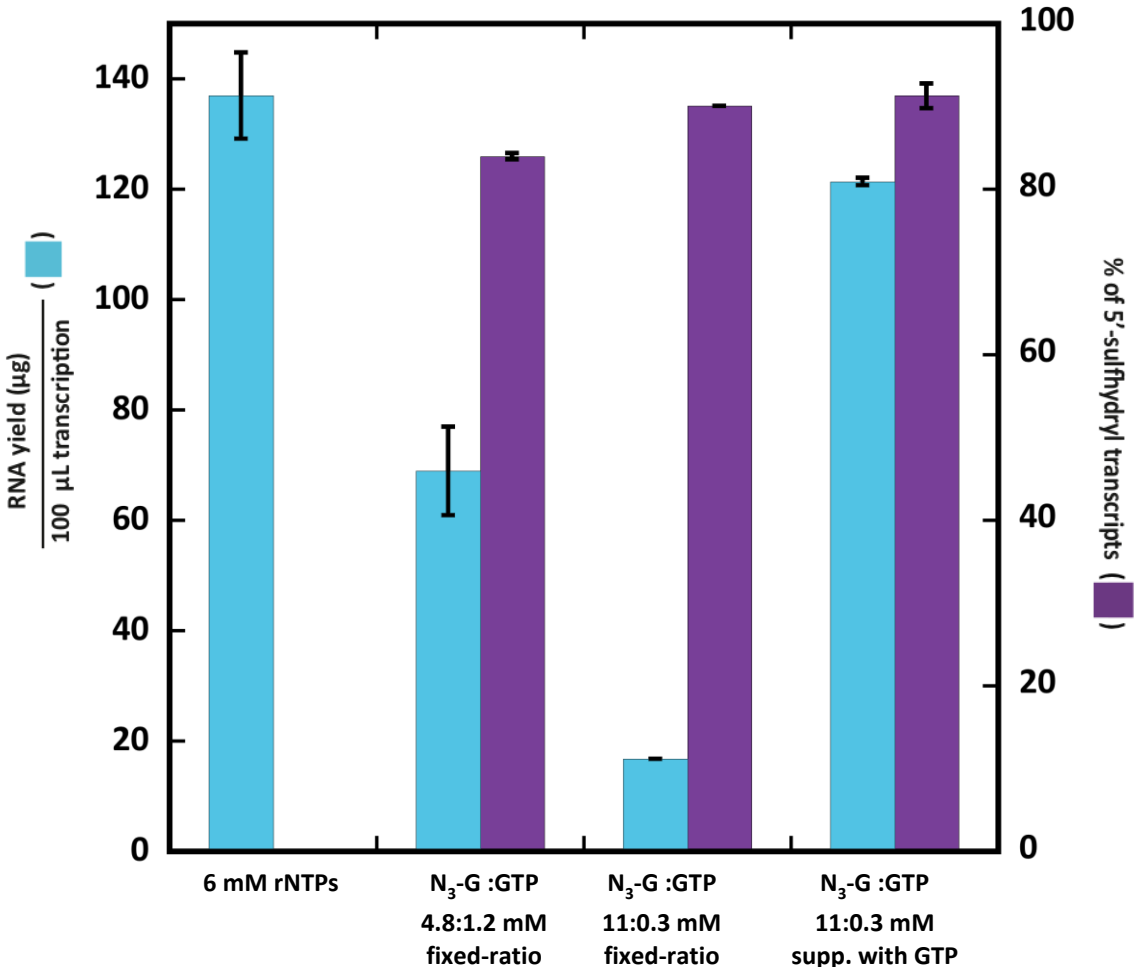


Figure 11. IVT using N_3-G as the transcriptional initiator: results from fixed-ratio and phased-addition approaches. The blue bars represent the yield of RNA in µg from a 100 µL IVT and the purple bars represent the percent of transcripts with a 5'-azido substitution.

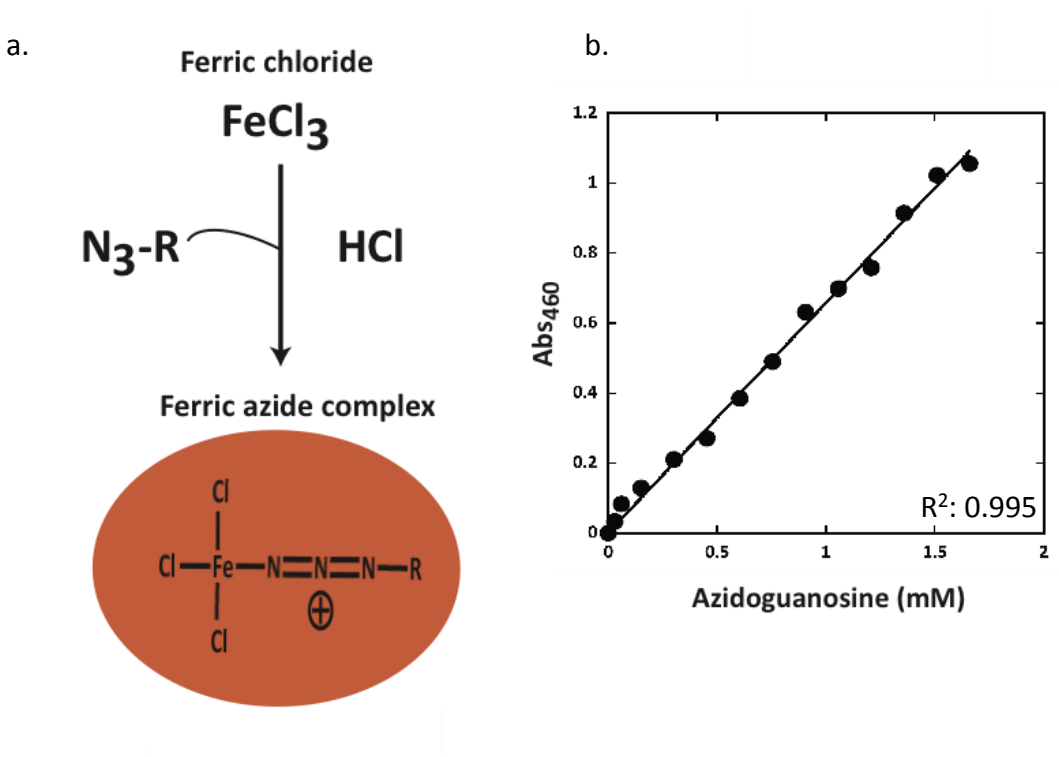


Figure 12. Quantitation of 5'-azido incorporation in an RNA. (a) Reaction of ferric chloride with an azide to produce a ferric azide complex that absorbs light at 460 nm. (b) Linear colorimetric standard curve with N_3-G.

Chapter 4
Discussion

4.1 Incorporation of 5'-deoxyguanosine-5'-monophosphorothioate and 5'-azido-5'-deoxyguanosine

It is vital to determine the amount of RNA modified with the guanosine analogs, because it provides the maximum amount of RNA that can be further labeled with a biochemical or biophysical probe. The extent of 5'-incorporation of all guanosine analogs was determined through incorporation spectrophotometric assays, resulting in a firm estimate of the amount of 5'-modified pre-tRNA^{Cys}.

I was able to incorporate GSMP at the 5'-terminus of pre-tRNA^{Cys} by using the phased-addition strategy in a run-off IVT using T7 RNA polymerase to generate 122 µg from a 100 µL IVT with 90% of the transcripts initiated with GSMP, as determined using the Ellman's test (Ellman, 1959). Similarly, I was able to incorporate N_3-G at the 5'-terminus of pre-tRNA^{Cys} by using the phased-addition strategy in a run-off IVT using T7 RNA polymerase to generate 116 µg in a 100 µL IVT with 92% of the transcripts initiated by N_3-G determined by use of the ferric chloride test (Christova-Bagdassarian and Atanassova, 2007; Tsuge et al., 2001).

In comparison to a typical IVT, I determined that there is a 2-fold (~52%) decrease in the yield of RNA when a 4:1 GSMP:GTP ratio is utilized. Additionally, I confirmed that by increasing the GSMP:GTP ratio, the GSMP could out-compete GTP at initiation and used predominantly for initiation (~92% of the transcripts). The phased-addition strategy also improved the yield such that there is only 11% decrease compared to a standard transcription. In the case of N_3-G, using 4:1 N_3-G:GTP ratio results in a 50% decrease in

RNA yield relative to the standard IVT; however, using the phased-addition strategy results in only a 16% decrease from a standard transcription. While tweaking the GSMP:GTP or N_3-G:GTP ratio and changing the number of GTP supplements might yield additional improvements, the results reported here already confirm that we can obtain 90% 5'-incorporation of the modified guanosine analog at ~90% of the typical IVT yield.

4.2 Incorporation of 5'-deoxy-5'-hydrazinylguanosine

I was able to incorporate NH_2NH-G at the 5'-terminus of pre-$tRNA^{Cys}$ by using the phased-addition strategy in a run-off IVT using T7 RNA polymerase to generate 105 µg from a 100 µL IVT with 62 % incorporation of NH_2NH-G, as determined using the veratraldehyde test (Kaveeshwar and Gupta, 1992). With use of the phased-addition strategy, I was able to improve the RNA yield by 11-fold compared to the fixed-ratio approach. Additionally the incorporation increased from 37% to 62%.

The NH_2NH-G incorporation results highlight an interesting dilemma, since T7 RNA polymerase does not appear to favor NH_2NH-G over GTP. When I increased the ratio between NH_2NH-G and GTP, such that there was a high concentration (22 mM) of NH_2NH-G and only 0.5 mM of GTP, the incorporation was still low. However, when the ratio between NH_2NH-G and GTP was increased, while lowering the concentrations of NH_2NH-G and GTP, I determined that there was higher incorporation of NH_2NH-G. Clearly, a higher concentration of NH_2NH-G has an inhibitory effect, and results in a lower yield.

Again, while the results are promising, there is always room for fine-tuning the ratio between NH_2NH-G and GTP, as well as the amount of GTP added during transcription or the number of times GTP is added during an IVT.

4.3 Conclusions

The results from all three guanosine analogs show that the phased-addition strategy enhances the yield of RNAs obtained through IVT, while ensuring that the vast majority of transcripts are initiated with the guanosine analog and not GTP. Central to these findings was also the ability to use two spectrophotometric assays to determine the amount of 5'-hydrazinyl and 5'-azido-modified RNAs.

The use of dinucleotides with a 5'-modification (e.g., 5'-azido-GpG) is a possible solution to the low yield problem encountered with the fixed-ratio approach. However, the dinucleotides are rather expensive, ranging between around $320 and $750 for a 1 µmole scale synthesis (Bio-Synthesis, Inc.). Due to the cost, the phased-addition strategy is favorable because it employs affordable and easily prepared reagents.

The phased-addition strategy, demonstrated here for 5'-labeling, could also be implemented to produce an internal modification by exploiting the DNA splint-based RNA ligation method (Moore and Sharp, 1992). Consider an RNA of 100 nucleotides in which modification is desired at position 66. Two different *in vitro* transcriptions can be preformed to generate two RNAs corresponding to positions 1 – 65 and 66 – 100. The shorter RNA (66 – 100, 35 nt) could be initiated using a 2’-modified guanosine 5’-monophosphate (or adenosine 5’-monophosphate, depending on the promoter of T7

RNA polymerase). The transcription with the modified guanosine 5’-monophosphate could be carried out using the phased-addition strategy, which will insert the 2'-modifed guanosine at the 5'-terminus of the transcript. The 5'-phosphate of this shorter transcript could then be ligated to the longer RNA (1 – 65, 65 nt) using a DNA splint and DNA ligase (Figure 13).

The first step in site-specific modification of RNAs is to generate RNAs with the appropriate reactive handles and in large amounts. The phased-addition strategy helps fulfill this objective. An important future direction will be to validate the phased-addition strategy with different RNAs of varying compositional bias and for internal labeling.

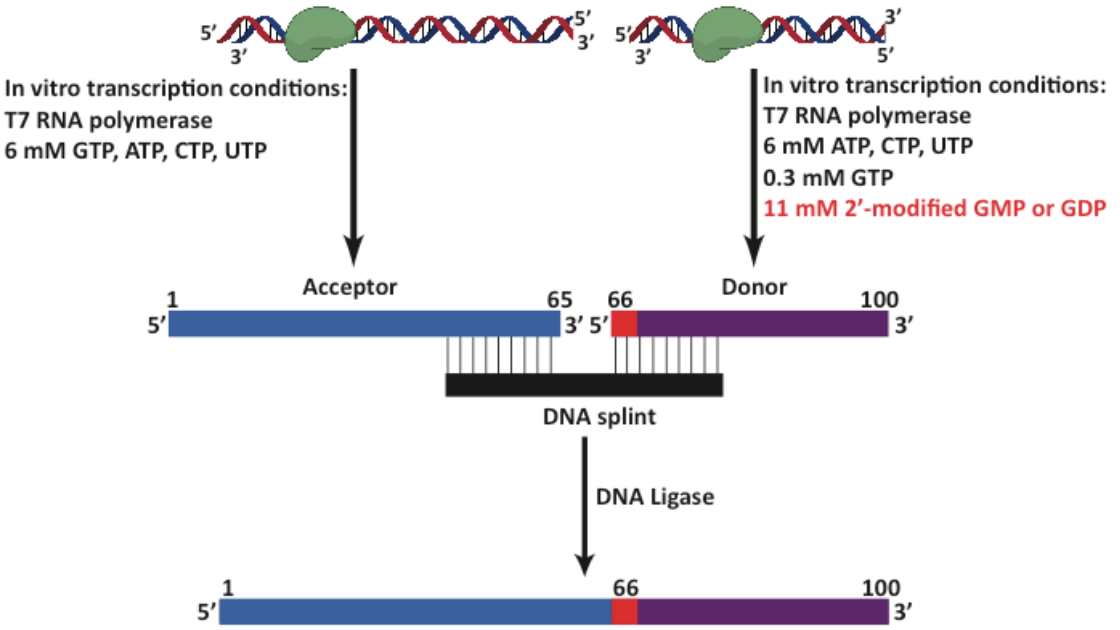


Figure 13. Application of the phased-addition strategy for introducing an internal modification. While the DNA splint-based RNA ligation approach of Moore and Sharp (1992) is key to generating the desired RNA with an internal site-specific modification, the phased-addition strategy could enhance the overall yield by ensuring that the 3′-fragment is predominantly 5′-modified.

References

Behrman EJ (2000) An improved synthesis of gunaoinse 5'-monothiophosphate. *J. Chem. Research* 446-447.

Blakeley BD, DePorter SM, Mohan U, Burai R, Tolbert BS, McNaughton BR (2012) Methods for identifying and characterizing interactions involving RNA. *Tetrahedron* 68: 8837-8855.

Brear P, Freeman GR, Shankey MC, Trmčić M, Hodgson DRW (2009) Aqueous methods for the preparation of 5'-substituted guanosine derivatives. *Chem. Comm.* 33: 4980-4981.

Cech TR (2012) The RNA worlds in context. *Cold Spring Harb. Perspect. Biol.* 4: a006742.

Christova-Bagdassarian V, Atanassova M (2007) Spectrophotometric determination of sodium azide in workplace air. *J. Univ. Chem. Technol. Metallurgy* 42: 311-314.

Domick C, Eggert F, Kath-Schorr S (2015) Site-specific enzymatic introduction of a norbornene modified unnatural base into RNA and application in post-transcriptional labeling. *Chem. Commun.* DOI: 10.1039/C5CC01765C

Ellman GL (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 82: 70-77.

Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35: 849–857.

Kaveeshwar R, Gupta VK (1992) A new spectrophotometric method for the determination of hydrazine in environmental samples. *Anal. Chem.* 344: 114-117.

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschiling DE, Cech TR (1982) Self-splicing RNA: autoexcision and autocyclization of ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* 31: 147-157.

Naga Prashant K, Anupama S (2014) Synthesis and biological evaluation of some novel Schiff's bases of veratraldehyde. *Asian J. Pharm. Clin. Res.* 2: 65-68.

Pagano JM, Clingman CC, Ryder SP (2011) Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* 17: 14-20.

Paredas E, Das SR (2011) Click chemistry for rapid labeling and ligation of RNA. *Chem. Biochem.* 12: 125-131.

Paredas E, Evans M, Das SR (2011) RNA labeling, conjugation and ligation. *Methods* 54: 251-259.

Qin PZ, Pyle AM (1999) Site-specific labeling of RNA with fluorophores and other structural probes. *Methods* 18: 60-70.

Raddatz S, Mueller-Ibeler J, Kluge J, WäB L, Burdinski G, Haven JR, Onofrey TJ, Wang D, Schweitzer M (2002) Hydrazide oligonucleotides: new chemical modification for chip array attachment and conjugation. *Nucleic Acids Res.* 30: 4793-4802.

Royer CA, Scarlata SF (2008) Fluorescence approaches to quantifying biomolecular interactions. *Methods Enzymol.* 450: 79-106.

Schmitz AG, Zelger-Paulus S, Gasser G, Sigel RKO (2014) Internal labeling strategy of large RNAs with minimal perturbation using fluorescent PNA. *ChemBioChem.* DOI: 10.1002/cbic.201500180

Skipsey M, Hack G, Hooper TA, Shankey MC, Conway LP, Schröder M, Hodgson DRW (2013) 5'-deoxy-5'-hydrazinylguanosine as an initiator of T7 RNA polymerase-catalyzed transcriptions for the preparation of labeling-ready RNAs. *Nucleos. Nucleot. Nucl.* 32: 670-681.

Tor Y, Dervan PB (1993) Site-specific enzymatic incorporation of an unnatural base N6-(6-aminohexyl)isoguanosine, into RNA. *J. Am. Chem. Soc.* 115: 4461-4467.

Tsuge K, Kataoka M, Seto Y (2001) Rapid determination of cyanide and azide in beverages by microdiffusion spectrophotometric method. *J. Anal. Toxicology* 25: 228-236.

Wallace A (2013) Fluor-labeling of RNA and fluorescence-based studies of precursor-tRNA cleavage by *Escherichia coli* ribonuclease P. *MS thesis*, The Ohio State University, Columbus, OH.

Zhang B, Cui Z, Sun L (2000) Synthesis of 5'-deoxy-5'-thioguanosine-5'-monophosphorothioate and its incorporation into RNA 5'-termini. *Org. Lett.* 3: 275-278.

Zhang L, Sun L, Zhiyong C, Gottlieb RL, Zhang B (2001) 5'-Sulfhydryl-modified RNA: Initiator synthesis, *in vitro* transcription, and enzymatic incorporation. *Bioconj. Chem.* 12: 939-948.